NASA-CR-199136

PSB NDB
7N-71-CR
63092
P. 9

On the role of head-related transfer function spectral notches in the judgement of sound source elevation

Ewan A. Macpherson
Waisman Center
University of Wisconsin-Madison
1500 Highland Avenue
Madison, WI 53705-2280
macpherson@waisman.wisc.edu

Abstract

Using a simple model of sound source elevation judgement, an attempt was made to predict two aspects of listeners' localization behavior from measurements of the positions of the primary high-frequency notch in their head-related transfer functions. These characteristics were: 1) the scatter in elevation judgements, and 2) possible biases in perceived elevation introduced by front-back and back-front reversals. Although significant differences were found among the notch-frequency patterns for individual subjects, the model was not capable of predicting differences in judgement behavior. This suggests that a simple model of elevation perception based on a single spectral notch frequency is inadequate.

1 Introduction

The role of spectral cues in auditory localization is known to be significant but is as yet poorly understood. While it has been established that low-frequency interaural time difference is the primary determinant of the left-right component of perceived source position [1], no simple and reliable cue for the elevation or front-back components has been found.[1] There does exist a regular dependence of spectral notch frequency on position for the head-related transfer functions of the cat [2,3] and somewhat similar feature motion for humans [4], and some researchers have proposed that this may be the principal elevation cue [5]. Although notch frequency clearly depends on position, it may be the case that the pattern is not as regular for humans as it is for the cat, and no causal relationship between this aspect of the physical acoustics and listener behavior has been confirmed.

The aims of the present study were to examine the differences among the notch frequency patterns of a number of individuals and to attempt to predict patterns in their elevation judgements on the basis of these differences. Predictions were made using the following simple model of elevation perception, which will be referred to as the single-notch model:

[1]The position of a source in space can be defined in a three-pole coordinate system with dimensions of left-right, back-front and elevation (up-down). The left-right dimension corresponds to the angle between the source and the vertical median plane. Sources with equal left-rightness lie on a "cone of confusion", so-called because the interaural time-difference cue is approximately constant and hence ambiguous

(NASA-CR-199136) ON THE ROLE OF HEAD-RELATED TRANSFER FUNCTION SPECTRAL NOTCHES IN THE JUDGEMENT OF SOUND SOURCE ELEVATION (Wisconsin Univ.) 9 p

N95-71576

Unclas

Z9/71 0063092

> Given that a source is localized to a particular cone of confusion (determined by interaural time difference) and to either the front or rear hemisphere (determined by some unknown spectral cue), then perceived elevation is determined by the frequency of the primary high-frequency notch in the head-related transfer function of the ear nearest the source.

Whatever plausibility of this model possesses rests on the observation that contours of equal notch frequency tend to intersect each cone of confusion only twice - once in the frontal hemisphere and once in the rear. This is generally true for moderate positive and negative elevations. Musicant and Butler [6] established that spectral features due to the filtering by the near ear are the dominant cues for resolving source position on the cone of confusion. Observations made in our laboratory and by Morimoto and Aokata [7] confirm that listeners are accurate in determining on which cone of confusion a source lies and that errors are primarily made in resolving position on the cone.

Using this model and measured notch patterns, two predictions pertaining to listeners' localization judgements were made. The first concerned the variance of the elevation responses and the second response bias under conditions of front-to-back or back-to-front confusion. The predictions were evaluated using free-field localization response data.

2 Methods

2.1 Subjects

Data were collected from six members of the Hearing Development Research Laboratory subject pool. There were three female and three male subjects ranging in age from 20 to 24. All reported normal hearing. For each subject head-related transfer functions were measured and free-field localization judgement data were collected as described below.

2.2 HRTF notch measurements

The procedure used to measure head-related transfer functions is described in detail by Wightman and Kistler [8]. Using probe-tube microphones positioned as close to the eardrum as possible, source-to-eardrum impulse responses were measured for positions spaced by 10° in both azimuth and elevation.

The location of the primary high-frequency spectral notch in each transfer function was located "by eye" on a computer screen plot of the spectrum and was marked using a mouse input device. Some judgement was required to select the desired notch; care was taken to follow particular features to higher elevations where they tended to peter out. The primary notch is visible in Figure 1, which shows directional transfer functions (HRTFs normalized by the diffuse-field response) as a function of elevation at 0° azimuth for subject SNF. Note the motion of the high-frequency notches as elevation increases. Since the extraction was a time-consuming task, the analysis was limited to positions spaced by 30° in azimuth and to elevations lying between -50° and +50°. This was done for both left and right ears and resulted in 264 data points for each subject.

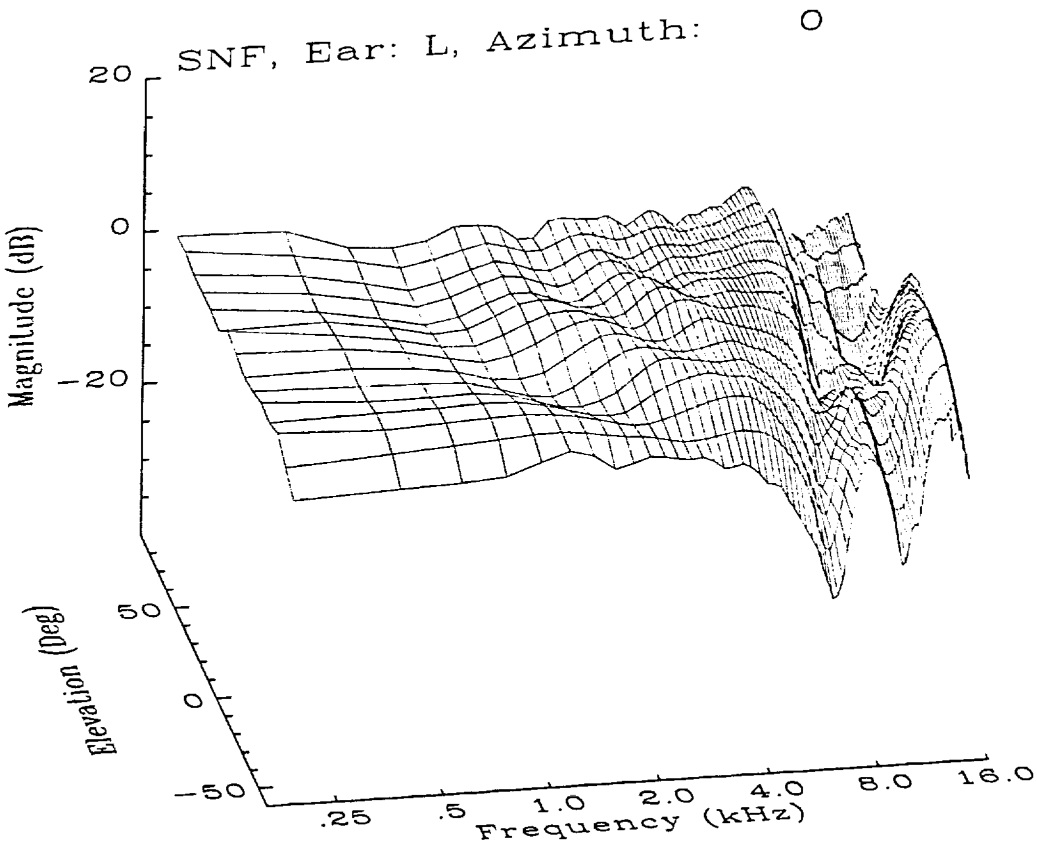

Figure 1: Directional transfer functions as a function of elevation at 0° azimuth for subject SNF.

2.3 Free-field judgements

Free-field localization judgements were collected with participants seated blindfolded in an anechoic chamber. Broadband (200-14000 Hz) noise bursts of 250 ms duration were played from loudspeakers mounted on a moveable arc. Subjects responded verbally with the azimuth and elevation of the perceived source location.

3 Individual notch frequency patterns

Contour plots of left-ear primary notch frequency as a function of direction are plotted in Figures 2-5 for four representative subjects. The dotted curves in these plots show the cones of confusion. Subjects SNF and SNX show approximately horizontal orientation of the notch contours on the ipsilateral side (negative azimuths). The contours for SNF are generally more closely spaced than those for SNX, revealing that notch frequency varies more slowly with position for the latter. Subjects SNT and SNY show upwards tilting of the contours towards the front. This is extreme in the case of SNY.

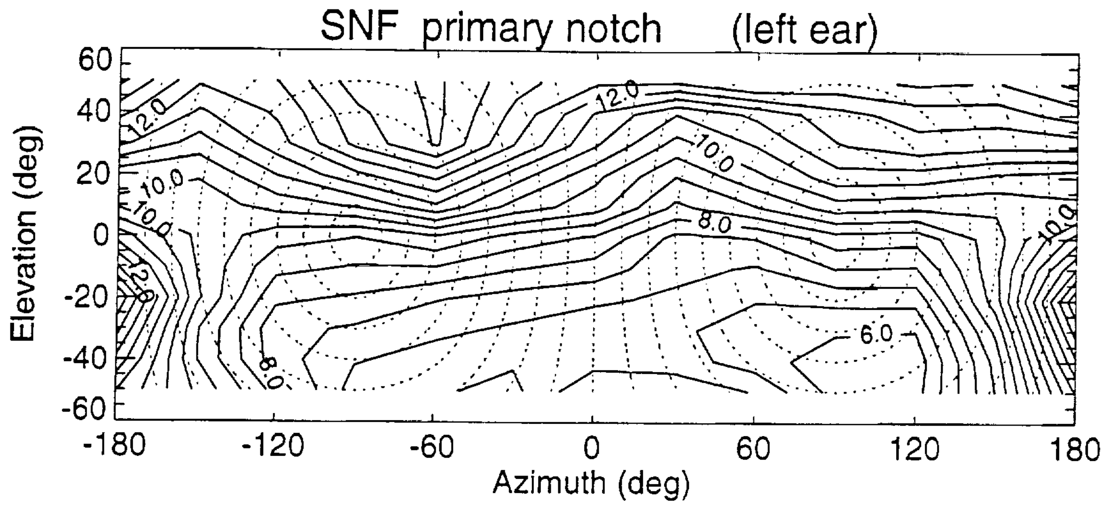

Figure 2: Contours of equal notch frequency (in kHz) for subject SNF. Dotted curves indicate cones of confusion.

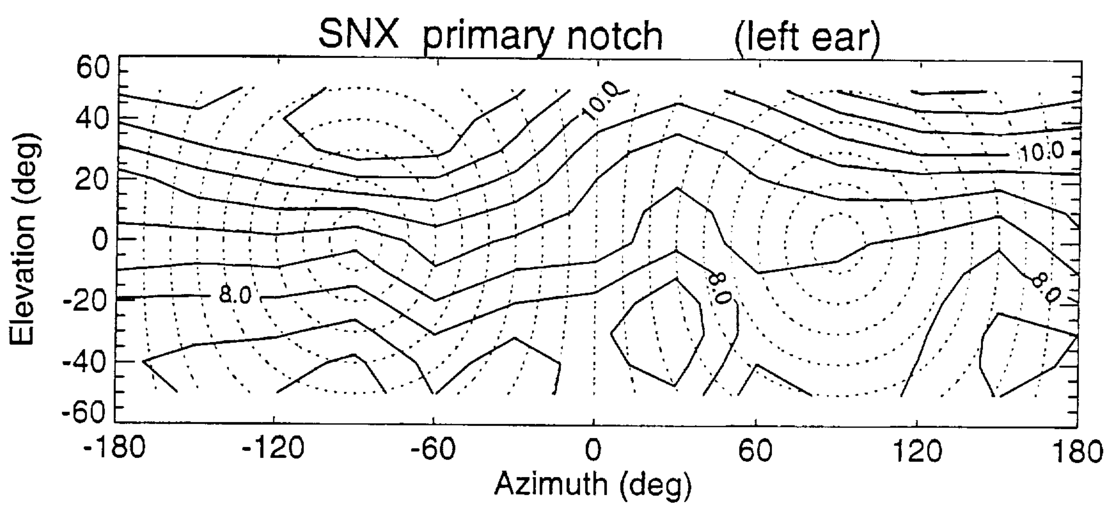

Figure 3: Contours of equal notch frequency for subject SNX.

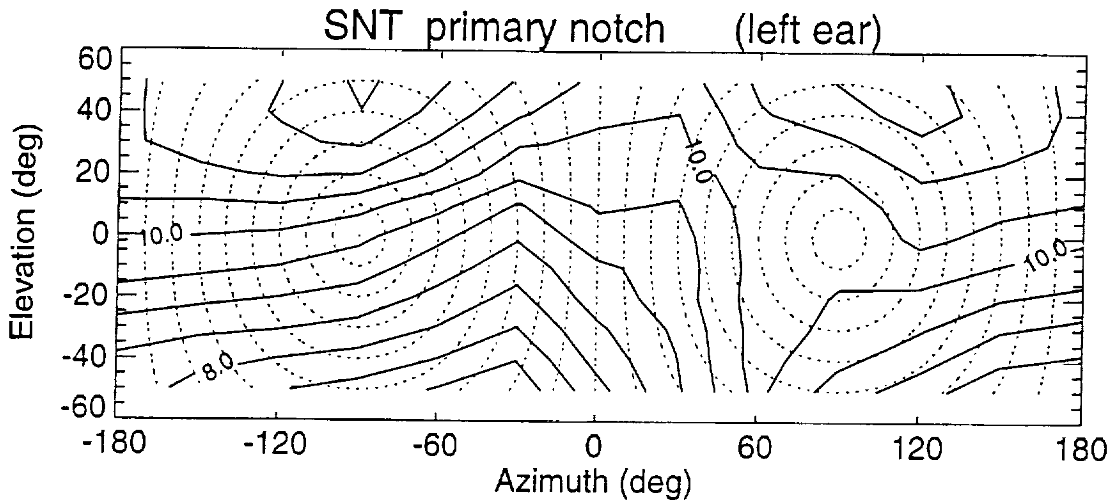

Figure 4: Contours of equal notch frequency for subject SNT.

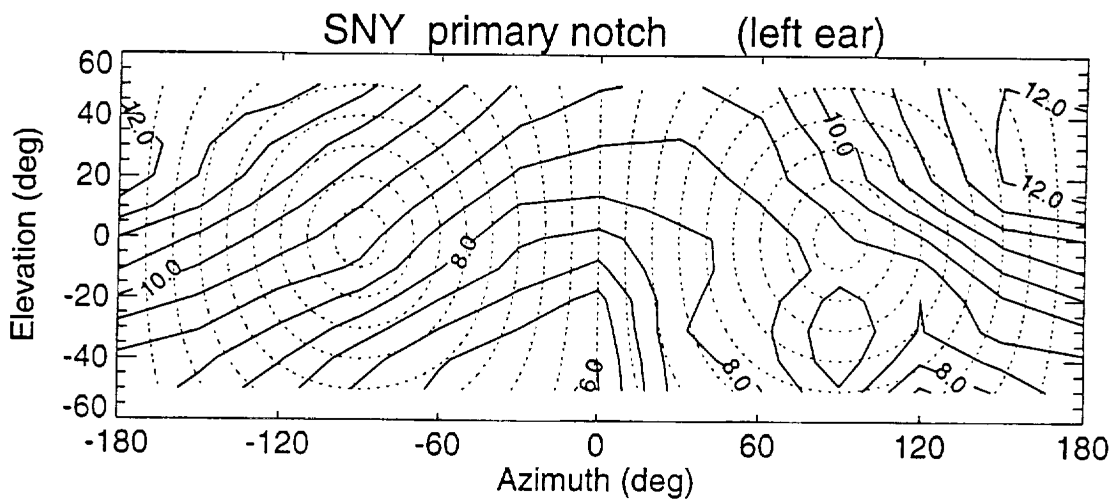

Figure 5: Contours of equal notch frequency for subject SNY.

4 Scatter of elevation judgements

4.1 Single-notch model prediction

The first prediction generated using the single-notch model concerns the relationship between the rate at which notches change frequency with position and the variance of the elevation component of subject's responses. If notch frequency determines elevation, then subjects for whom notches move more rapidly should show less spatial scatter in their responses to individual real source locations. Uncertainty about notch frequency should correspond to relatively greater uncertainty about elevation for subjects with "slow" notches.

4.2 Data analysis

To evaluate this prediction, some measure of the spatial dependance of notch frequency for each subject was required. The magnitude of the near-ear notch frequency gradient averaged over the region of the sphere under consideration was chosen as a suitable metric of overall notch "speed". This value (∇) was calculated over all available positions and also for positions within the region of the upper hemisphere lying between -30° and +30° (the high-front case). The latter region was considered to be of particular interest since, unlike in the coronal plane, elevation judgement is likely to be almost entirely spectrally-based near the median plane due to the near-zero values of interaural difference cues.

To characterize the degree of scatter in subjects' judgements, the standard deviation of the elevation responses elicited by each physical source position was calculated and then averaged over the region of interest, yielding the value σ. Only responses classified as unconfused were analyzed; those deemed to be examples of front-back, back-front, or up-down reversal were excluded. The number of responses remaining at each location ranged from 4 to 7. The criteria for these classifications are discussed in Section 5.2.

4.3 Results

The results of these analyses are presented in Table 1. Linear regression revealed the correlations between ∇ and σ to be -0.21 in the overall case and -0.61 in the high-front case.

5 Bias in Front-Back Confused Elevation Judgements

5.1 Single-notch model prediction

The second prediction made by the single-notch model concerns the effects of front-to-back and back-to-front confusions on elevation error. If, as in the cases of subjects SNT and SNY, the contours of constant notch frequency are tilted significantly away from the horizontal, and if the single-notch model is correct, then a front-back or back-front reversal should have a significant and consistent effect on elevation judgement errors. For example, it might be expected that SNY would experience over-elevation in cases of back-to-front reversal since the notch contours sweep upwards towards the front. Similarly, front-to-back reversals should be under-elevated.

Table 1. Averaged Standard Deviation of Elevation Judgements

Subject	All Positions		High-Front	
	∇ (Hz/deg)	σ (deg)	∇ (Hz/deg)	σ (deg)
SNF	67.1	13.0	91.6	11.3
SNJ	61.9	15.5	59.6	16.0
SNR	43.5	16.5	42.5	21.4
SNT	47.0	16.6	37.2	13.2
SNX	40.1	11.8	36.3	17.4
SNY	39.1	16.3	22.9	17.8
	correlation = -0.21		correlation = -0.61	

5.2 Data analysis

The available responses for each physical source location in the left hemisphere were classified as one of: correct front (F), correct back (B), front-to-back reversed (FB), back-to-front reversed (BF), or up-down reversed (UD). If a judged position lay closer to the real location when reflected in the coronal plane, it was deemed to be a BF or FB confusion. Errors of elevation of greater than 90° were classed as up-down confusions and were excluded from the analysis. The mean difference between the reported and actual elevations was calculated at each position, and then these mean differences were averaged over the region of interest. This procedure was carried out for subjects SNF, SNX, SNT, and SNY, who all had orderly notch patterns and made significant numbers of front-back reversals.

5.3 Results

The results are presented in Table 2, in which arrows in cells indicate the predicted direction of the bias. Both of the subjects with more horizontal contours tend to over-elevate, although their bias patterns differ. Subject SNT does show significant over-elevation of back-to-front reversed judgements, but also over-elevates sources correctly localized in the front. SNY, for whom the notches were even more strongly tilted, shows no significant bias in any condition. The striking result is that for both of these listeners there are no differences between confused and unconfused judgements.

Table 2. Elevation Judgement Bias (bias in degrees).

Response Type	Subjects with Horizontal Contours		Subjects with Tilted Contours	
	SNF	SNX	SNT	SNY
F	12.1	10.0	26.7	0.7
BF	17.1	1.1	↑ 24.7	↑ 2.7
B	8.4	6.1	10.3	-4.3
FB	-1.2	5.0	↓ 11.2	↓ 2.1

6 Discussion and conclusions

Although significant differences exist among the notch patterns for different subjects, attempting to predict localization behavior on the basis of these differences using the single-notch model cannot be termed a success. There appears to be no strong relationship between the average magnitude of the notch frequency gradient and response scatter either for the all-positions or the high-front case. Although the correlation coefficient of -0.61 is suggestive it is not a convincing result, and its magnitude is due mainly to one outlying point (subject SNF). There appears to be little evidence of a relationship when the quantities are averaged over all positions. It is not surprising that the observed correlations, while low, were in the appropriate direction since the rate of notch movement with position must be positively correlated with the rate of change of overall spectral shape with position.

In the case of front-back and back-front reversals, the predictions of the simple notch model were not observed. The two subjects (SNT and SNY) with tilted notch contours had very different error bias patterns and, more importantly, showed no effect of front-back reversals on elevation judgements.

The results of these analyses clearly do not support the single-notch model of elevation perception. The observed individual differences in notch-frequency variation do not yield strong predictive power for localization behavior when coupled with this model. Therefore, elevation judgements must depend on additional spectral cues which have yet to be identified and verified.

Acknowledgements

The author is grateful to Dr. Frederic Wightman and Dr. Doris Kistler for their advice and assistance. The tedious extraction of notch-frequency from the HRTF plots was done by Kristin Andersen. This research was supported in part by grants from NIH, NASA, and ONR.

References

[1] Wightman, F.L. and D.J. Kistler. "The Dominant Role of Low-Frequency Interaural Time Differences in Sound Localization." *Journal of the Acoustical Society of America* **91** (1992): 1648-1661.

[2] Rice, J.J., B.J. May, G.A. Spirou, and E.D. Young. "Pinna-Based Spectral Cues for Sound Localization in Cat." *Hearing Research* **58** (1992): 132-152.

[3] Neti, C., E.D. Young, and M.H. Schneider. "Neural Network Models of Sound Localization Based on Directional Filtering by the Pinna." *Journal of the Acoustical Society of America* **92** (1992): 3140-3156.

[4] Kuhn, G.F. "Physical Acoustics and Measurements Pertaining to Directional Hearing." In *Directional Hearing*, edited by W.A. Yost and G. Gourevitch, 3-25. New York: Springer-Verlag, 1987.

[5] Butler, R.A. and K. Belendiuk. "Spectral Cues Utilized in the Localization of Sound in the Median Sagittal Plane." *Journal of the Acoustical Society of America* **61** (1977): 1264-1269.

[6] Musicant, A.D. and R.A. Butler. "The Influence of Pinnae-Based Spectral Cues on Sound Localization." *Journal of the Acoustical Society of America* **75** (1984): 1195-1200.

[7] Morimoto, M. and H. Aokata. "Localization Cues of Sound Sources in the Upper Hemisphere." *Journal of the Acoustical Society of Japan* **5** (1984): 165-173.

[8] Wightman, F.L. and D.J. Kistler. "Headphone Simulation of Free-Field Listening I: Stimulus Synthesis." *Journal of the Acoustical Society of America* **85** (1989): 858-867.